1986-87 | E243

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications : 1986-87

FINAL EXAMINATION

PART - II

Statistical Inference (Practical)

Date: 22.6.1987 Maximum Marks: 100 Time: 2 hours

Note: You may answer any part of any question.

1. An investigator was asked to take 10 independent measurements on the maximum internal diameter of a pot. The standard deviation of these measurements was 0.0345 mm. The experiment was repeated after a few days when the investigator had sufficient practice, and the standard deviation of the 10 measurements was 0.0126 mm. Does this mean that with practice the investigator has become more consistent ? (Take $\alpha = .01$) [30]

2. The heart weights in gms. of 12 female and 15 male cats are given below. Does the heart of a male cat on an average weigh more than that of a female cat ?

Males	12.7,	15.6,	9.1,	12.8,	8.3,	11.2,	9.4,	8.0,
	14.9,	10.7,	13.6,	9.6,	11.7,	9.3,	7.6	
Females	7.4,	7.3,	7.1,	9.0,	7.6,	9.5,	10.1,	
	10.2,	10.1,	9.5,	8.7,	7.2			

(Take $\alpha = .01$) [30]

3. Let $x_1, \ldots, x_n$ are iid observations from the following density :

$$f_\theta(x) = \frac{1}{\theta} \quad 0 < x < \theta$$

Consider the critical region $x_{(n)} > 0.8$ for testing $H_0 : \theta = 1$, $H_1 : \theta \neq 1$ where $x_{(n)}$ is the largest of $x_1, \ldots, x_n$. What is the associated probability of error I and what is the power function ? [40]

p.t.o.

4. An urn contains white and black balls in unknown proportions, the total number of balls being 8. Three balls are taken at random, of which 2 are found to be white and 1 black. Find the maximum-likelihood estimate of the number of white balls in the urn.

[20

:bcc:

1986-87	E263

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications: 1986-87

FINAL EXAMINATION

PART - II

Industrial Statistics (Theory and Practical)

Date: 19.6.1987 Maximum Marks: 100 Time: 2 hours

Note: Answer FOUR questions altogether, choosing any two from each group.

GROUP A

1.(a) How does a control chart serve the purpose of a statistical test of hypothesis ?

(b) Discuss Type I and Type II errors relative to or control chart.

(c) Explain how the 'Control Chart analysis' of a set of observations on some variable characteristics may be considered equivalent to 'Analysis of variance in one way classification'.

(d) What is an acceptance sampling plan ?

(e) What are the basic differences in the objectives of control chart and acceptance sampling in quality control ?

(f) Explain the difference between the 'Control limits' and the 'natural process limits'.

(5+5+10+2+5+3) = [30]

2.(a) Mr. Bose took samples of size 100 for 10 days. The fraction defectives observed on each day were,

0.04, 0.08, 0.10, 0.03, 0.05, 0.07,
0.09, 0.11, 0.01 and 0.07.

Determine if the fraction defective of the process is in control.

Contd..... 2/-

Contd..... Q.No.2

(b) When Mr. Bose was on leave for a week, Mr. Anand took samples each day but the size varied. The sample sizes and the corresponding number of defectives observed were as follows:

day	sample size	no. of defectives
Monday	75	6
Tuesday	115	5
Wednesday	75	11
Thursday	50	10
Friday	115	7

Is the process still in control with respect to the average fraction defective observed in (a) but accounting for the variation in sample size ?

(c) A control chart for the number of defects is maintained on a process with $n\bar{p} = 16$. A sample of size 100 taken each day was analysed.

(i) What is the probability that a shift in the process average $n\bar{p} = 20$ will be detected on the first day following the shift ?

(ii) What is the probability that the shift will be detected by the end of the third day or earlier ?

(10+10+10) = [30]

3.(a) Samples of size n = 5 are taken from a manufacturing process every hour. A quality characteristic is measured and $\bar{X}$ and R are computed for each sample. After 25 samples we have,

$$\Sigma \bar{X}_i = 662.50 \; ; \quad \Sigma R_i = 9.00$$

(i) Find the control limits for the $\bar{X}$ and R charts.

(ii) Assuming both charts exhibit control, if specifications are 26.40 ± 0.50, estimate the fraction non-conforming.

(iii) If the mean of the process were 26.40, what fraction non-conformance would result ?

Contd..... 3

Contd..... Q.No.3

(b) The following fraction defective control chart with $n = 100$ is used to control a process:

UCL = 0.0750
Center line = 0.0400
LCL = 0.0050

(i) Use the poisson approximation to the binomial to find the probability of a type I error.

(ii) Use the poisson approximation to find the probability of a type II error if the true process fraction defective is 0.0600.

(iii) Draw the OC curve of this control chart using poisson approximation.

(3 x 5) + (2+2+8+3) = [30]

GROUP B

4.(a) Define AOQL and explain how do you obtain it for a single sampling plan.

(b) Obtain a single sampling plan when N = 3000, process average 2% and LTPD = 7% from Dodge and Romig Sampling plan.

(c) Draw the OC curve of the plan.

(6+4+10) = [20]

5.(a) Find an appropriate single sampling plan for a lot size of 1500, and an LTPD of 5.0% with a consumer risk of 0.10 if the estimated process average is 0.6% defective.

(b) In acceptance sampling under MIL-STD 105D single sampling is to be used with inspection level II, an AQL of 4.0% and a lot size of 2500. What are the appropriate sampling plans under (i) Normal inspection and (ii) Tightened inspection ?

Contd..... 4/-

Contd..... Q.No.5

(c) What are the acceptance criteria under reduced inspection for the conditions stated in (b) above ?

(d) Assume that an 8% defective lot is submitted under the reduced inspection plan of (c). Use poisson approximation to compute approximate probabilities that (i) the lot will be accepted and reduced inspection continued, (ii) the lot will be rejected.

(3+6+3+8) = [20]

:bcc:

1986-87 | E263

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications : 1986-87

FINAL EXAMINATION

PART - II

Design of Experiments (Theory and Practical)

Date: 17.6.1987 Maximum Marks: 100 Time: 2 hrs.

Note: Answer any FOUR questions.

1.(a) What is the logic of experimentation of designs ?

(b) What are the requirements of a good experiment. Show how the above requirements are being realised in completely randomized design, randomized block design and latin square design.

(5+20) = [25]

2.(a) Write down important assumptions underlying analysis of variance models.

(b) Define main effects and interaction of 2^2 factorial experiment and show that they are mutually orthogonal treatment contrasts.

(10+15) = [25]

3.(a) Set up the analysis of variance table of a randomized block design with v treatments in b blocks.

(b) Denoting by t_i effect due to ith ($i = 1, \ldots, v$) treatment, how will you test the hypotheses (i) $t_1 = t_2$ (ii) $t_1 + t_3 = 2t_2$?

(c) How will you obtain confidence interval for $t_1 - t_2$.

(7+13+5) = [25]

4. The following table gives plan and yield of sugar cane (in a suitable unit) per plot. Analyse the data to find out if there are any treatment effects.

Contd..... 2/-

Contd..... Q.No.4

Row	Column				
	I	II	III	IV	V
	A	E	D	C	B
I	52.5	46.3	44.1	48.1	40.9
	D	B	A	E	C
II	44.2	42.9	51.3	49.3	32.6
	B	A	C	D	E
III	49.7	47.3	38.1	41.0	47.2
	C	D	E	B	A
IV	43.2	42.5	67.2	55.1	45.3
	E	C	B	A	D
V	47.0	43.2	46.7	46.0	43.2

[25]

5. The following table gives the plan and yields of a manurial experiment involving two factors A,B each at two levels. Analyse the data.

Block	Yields			
	(1)	a	b	ab
1	25.5	40.0	25.5	10.5
	a	b	ab	(1)
2	60.0	40.5	20.5	36.5
	b	a	(1)	ab
3	15.5	30.0	20.0	15.5
	a	(1)	ab	b
4	35.0	15.5	15.5	30.5

[25]

:bcc:

1986-87	E253

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications: 1986-87

FINAL EXAMINATION

PART - II

Sample Surveys (Practical)

Date: 15.6.1987 Maximum Marks: 100 Time: 2 hours

Note: The paper contains 130 marks. Answer as many as you can. Maximum you can score is 100.

1. In a sample of 50 households drawn with SRSWOR from a village consisting of 250 households, only 8 households were found to possess a bicycle. These had 3, 5, 3, 4, 7, 4, 4 and 5 members respectively. Estimate unbiasedly the total number of households in the village possessing a bicycle (H_6) as well as the total number of persons in such households (P_6). Also estimate the rse's of these estimates by using the unbiased estimates of their variances.

[15]

2. A list of 30 villages in a tehsil arranged in ascending order of geographical area (x) is given in the following table together with village-wise area under winter paddy(y)

No.	x	y	No.	x	y	No.	x	y
1	103	41	11	145	74	21	209	93
2	106	33	12	147	13	22	216	38
3	120	87	13	151	81	23	224	87
4	120	87	14	153	41	24	229	72
5	121	56	15	160	58	25	230	127
6	121	62	16	166	44	26	235	114
7	124	58	17	176	65	27	238	82
8	128	19	18	178	69	28	240	108
9	135	64	19	185	29	29	241	94
10	137	61	20	206	46	30	243	116

Contd..... 2/-

Contd..... Q.No.2

(i) Draw 3 circular systematic samples of 7 villages each.

(ii) Making use of the 21 sample observations obtained in (i) estimate the relative efficiency of circular systematic sampling as compared to that of srswor for estimating the total area under paddy (y) based on a sample of 7 villages.

(iii) Obtain a single combined estimate of Y based on all the 3 samples drawn in (i) and also estimate its relative standard error.

[25]

3. In a study of the possible use of sampling to cut down the worn in taking inventory in a stockroom, a count is made on the value of articles on each of 36 shelves in the room. The values to the nearest rupees are as follows:

29, 38, 42, 44, 45, 47, 51, 53, 53, 54, 56,
56, 56, 58, 58, 59, 60, 60, 60, 60, 61, 61,
61, 62, 64, 65, 65, 67, 67, 68, 69, 71, 74,
77, 82, 85.

The estimate of total value made from a sample is to be correct within Rs.200/- apart from a 1 in 20 chance. An advisor suggests that a simple random sample of 12 shelves will meet the requirement. Do you agree ?

$$\Sigma y = 2138, \qquad \Sigma y^2 = 131,682.$$

[10]

4. In a sample survey for estimating the number of standards of pepper in a tehsil having 72 villages, a sample of 6 villages was selected with Simple Random Sampling with Replacement and from each sample village 5 clusters of 20 fields each were drawn with Simple Random Sampling without Replacement. Data on number of clusters in the sample villages and on the number of standards in the sample clusters are given below.

Contd..... 3/-

Contd..... Q.No.4

Sample village	No. of clusters	Number of standards in sample cluster (in 10)				
		1	2	3	4	5
1	27	43	40	36	98	39
2	14	116	55	306	172	127
3	116	69	22	84	122	58
4	118	104	133	118	73	196
5	26	91	45	3	12	24
6	91	34	0	9	3	34

Estimate unbiasedly the total number of standards in the tehsil and obtain its relative standard error by estimating unbiasedly its variance. [20]

5. The following data show the stratification of all the farms in a country by farm size and the average and s.d. of acres of corn per farm (Y) in each stratum.

Farm-size	No. of farms (N_h)	Average Acres of corn ($\bar{Y}_h$)	Standard Deviation (S_h)
0 - 40	394	5.4	8.3
41 - 80	461	16.3	13.3
81 - 120	391	24.3	15.1
121 - 160	334	34.5	19.8
161 - 200	169	42.1	24.5
201 - 240	113	50.1	26.0
241 -	148	63.8	35.2

For a sample of 100 farms compute the sample sizes in each stratum under (i) proportional allocation. (ii) Optimum allocation. Compare the variances of the estimates of the population mean of Y under these allocations.

[25]

6. The following table gives the household sizes and total monthly income of all the 24 households in a village. Select 5 households with probability proportional to household size with replacement and from the dates on total monthly income of these 5 households, obtain estimates of

Contd..... 4/-

Contd..... Q.No.6

total monthly income for the whole village along with its standard error.

Serial no. of household	Household size	Total monthly income (in Rs.)
1	5	472
2	3	678
3	3	1039
4	7	1050
5	4	835
6	4	380
7	6	636
8	6	712
9	4	408
10	5	592
11	3	365
12	7	894
13	6	672
14	4	532
15	5	2039

[20]

7. From the population of problem 2 draw a PPSWR sample of 5 villages taking geographical area as the measure of size by Lahiris method. Obtain an unbiased estimate of total area under paddy and an unbiased estimate of its variance.

[25]

: bcc:

1986-87 | E253

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications: 1986-87

FINAL EXAMINATION

PART - II

Sample Surveys (Theory)

Date: 12.6.1987 Maximum Marks: 100 Time: 2 hours

Note: Answer any FOUR questions. Each question carry equal marks. Maximum you can score for each subdivision of a question has been indicated in the margin.

1.(a) Find an unbiased estimator of variance of sample mean under Simple Random Sampling without Replacement (SRSWOR).

(b) Consider estimator of the form $t = \sum_{i \in s} a_i Y_i$ where $a_1, \ldots, a_N$ are constants attached to the unity 1, 2,...,N respectively of the population and Y_i is the value of the characteristic y on the unit i in the population. Show that under SRSWOR, t is unbiased for $\bar{Y}$, iff $a_i = \frac{1}{N}$ $\forall$ $i = 1, \ldots, N$.

(c) Compare between SRSWOR and Simple Random Sampling with Replacement, both the schemes being based on n draws and the ordinary sample mean being used in both the cases for estimating the population mean. When is the one likely to be significantly preferable to the other from the statistical point of view ?

[10 + 8 + (4+3)] = [25]

2.(a) Under a linear cost function find optimum allocation of sample sizes for different strata when the objective is to minimise the variance of the estimator for a fixed total cost.

(b) With two strata, a samples would like to have $n_1 = n_2$ for administrative convenience, instead of using the values given by Neyman allocation. If $V(\bar{y}_{st})$, $V_{opt}(\bar{y}_{st})$ denote the variances given by the $n_1 = n_2$ and the Neyman

Contd..... 2/-

Contd..... Q.No.2.(b)

allocations, respectively, show that the fractional increase in variance

$$\frac{V(\bar{y}_{st}) - V_{opt}(\bar{y}_{st})}{V_{opt}(\bar{y}_{st})} = \left(\frac{r-1}{r+1}\right)^2$$

where $r = n_1/n_2$ as given by the Neyman allocation.

(13+12) = [25]

3.(a) Show that under Simple Random Sampling Without Replacement, $|\text{Bias of } \hat{Y}_R| \leq \text{c.v. } (\bar{x}) \times \text{S.E. of } \hat{Y}_R$ when $\hat{Y}_R$ denotes ratio estimator of population total and other symbols have their usual meanings.

(b) Obtain an expression for variance of an unbiased estimator of population mean under systematic sampling.

(13+12) = [25]

4.(a) What is a cluster sampling ? Assuming clusters of equal sizes obtain an unbiased estimator of population mean under cluster sampling. Obtain expression for its variance in terms of population intraclass correlation coefficient of units belonging to the same cluster. Hence obtain the efficiency of cluster sampling with respect to a comparable simple Random Sampling plan.

(b) Under PPSWR - scheme of sampling obtain the inclusion - probability of ith unit in the sample.

(4+2+10+4+5) = [25

5. Under Probability Proportional to Size With Replacement Sampling Design obtain an unbiased estimator of population total, its variance and an unbiased estimator of this variance. How would you select a sample by Lahiri's method ? Show that under Lahiri's scheme of sampling probability of selecting any unit at any particular draw is proportional to its size.

(4+5+5+5+6) = [25]

6. Obtain an unbiased estimator of population total under a Two Stage Sampling Design with SRSWOR at both the stages. Obtain expression for its variance. Obtain an unbiased estimator of the variance when sampling is by (a) SRSWR at first stage (b) PPSWR at first stage.

(4+12+3+6) = [25]

: bcc:

1986-87	E243

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications: 1986-87

PART - II

FINAL EXAMINATION

Statistical Inference (Theory)

Date: ~~10~~ 8.6.1987 Maximum Marks: 100 Time: 2 hrs.

Note: You may answer any part of any question.

1.(a) Define sufficient statistic. State Neyman - Fisher factorization theorem.

(5+5) = [10]

(b) Let $X_1, \ldots, X_n$ be iid and the distribution under θ of X_1 be the uniform over $[0, \theta]$ where $\theta > 0$ is unknown. Find a real-valued sufficient statistic for θ and show that it is sufficient. [10]

2.(a) State and prove the Rao-Blackwell theorem.

(5+15) = [20]

(b) Let $X_1, \ldots, X_n$ be iid and $P_\theta(X_1 = 1) = 1 - P_\theta(X_1 = 0) = \theta$ where $0 < \theta < 1$ is unknown. Find the minimum variance unbiased estimator T of 2θ; what is its variance ? Find the minimum variance unbiased estimator of the parametric function $\mathrm{var}_\theta(T)$.

(10+7+10) = [27]

3.(a) Define a maximum likelihood estimator. [5]

(b) Let $X_1, \ldots, X_n$ be as in question no.2.(b); find a maximum likelihood estimator $\hat{\theta}_n$ of θ ; in what sense it is unique ? Show that $\{\hat{\theta}_n\}$ is a consistent sequence of estimators of θ.

(15+5+10) = [30]

p.t.o.

4.(a) Let T be an unbiased estimator of $g(\theta)$. Is it true $f(T)$ is an unbiased estimator of $f(g(\theta))$ where f is a real-valued continuous function free from θ ? Justify your answer.

(2+6) = [8]

(b) Let $\{T_n\}$ be a consistent sequence of estimators of $g(\theta)$. Is it true that $\{f(T_n)\}$ is a consistent sequence of estimators of $f(g(\theta))$ where f is a real-valued continuous function free from θ ? Justify your answer.

(2+6) = [8]

5. Write a short note on the method of moments. [7]

:bcc:

1986-87	E263

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications : 1986-87

PERIODICAL EXAMINATION

PART - II

Industrial Statistics (Theory and Practical)

Date: 27.5.1987 Maximum Marks: 100 Time: 2 hours

Note: (a) Answer any FOUR questions with at least one question from each group.

(b) All questions carry equal marks, figures in the margin indicate full marks with breakup for each part.

Group A

1.(a) When do you consider a process to be under statistical control ?

(b) Surface defects have been counted on 25 rectangular steel plates and the data are shown below. Set up a control chart for non-conformities using these data. Does the process producing the plates appear to be in statistical control ? If not find the standard and the revised control limits.

Plate No.	No. of defects	Plate No.	No. of defects	Plate No.	No. of defects
1	1	9	2	17	3
2	0	10	1	18	5
3	4	11	0	19	4
4	3	12	1	20	6
5	1	13	8	21	3
6	2	14	0	22	1
7	5	15	2	23	0
8	0	16	1	24	2
				25	4

Contd..... 2/-

Contd..... Q.No.1

(c) A control chart is to be established on a process producing refrigrators. The inspection unit is one refrigerator and a control chart for number of defects is to be used. As preliminary data, 360 defects were observed in inspecting 30 refrigerators.

(i) What are the 3-sigma control limits ?

(ii) What is the α-risk for this control chart ?

(iii) What is the β-risk if the average number of defects were actually 6 (i.e., if $\bar{c} = 6.0$) ?

(5+(6+4)+(2+4+4)) = [25]

2. The data below gives the number of defective bearing and seal assembly in samples of size 150.

Sample No.	No. of defective	Sample No.	No. of defective	Sample No.	No. o defec
1	7	8	5	15	5
2	4	9	2	16	1
3	1	10	7	17	4
4	3	11	6	18	5
5	6	12	15	19	7
6	8	13	0	20	12
7	10	14	9		

(a) Construct a number defective control chart for this data. If any point plot out of control, assume that assignable causes can be found and defermine the revised control limits.

(b) It is decided to take a sample of size 64 for future contro Obtain the appropriate limits and the central line for control chart for future production assuming that the process level remains unchanged.

(c) With respect to the control chart for future production with a sample size of 64, to what level must the fraction defective increase to make the β-risk equal to 0.50 ? Assume poisson distribution.

(6+4)+5+10) = [25]

3. The following data were collected from a process manufacturing power supply equipment. The variable of interest is output voltage and n = 5.

Sample No.	Output voltage		Sample No.	Output voltage	
	Average	Range		Average	Rang.
1	103	4	6	106	3
2	102	5	7	102	7
3	104	2	8	105	2
4	105	11	9	106	4
5	104	4	10	104	3

(a) Compute central line and control limits suitable for controlling future production and draw the control cha...

(b) Estimate the process standard deviation assuming that output voltage is normally distributed.

(c) Obtain the process capability (Natural tolerence) of the process.

(d) What would be your estimate of the process fraction non-conformance if the specifications are 103 ± 4 ?

(e) What approaches for reduction of the fraction non-conforming can you suggest ?

(f) Estimate the fraction non-conforming with your suggested approach.

((6+4)+2+2+5+1+5) = [25]

4. Control chart for $\bar{X}$ and R are maintained for a certain quality characteristic. The sample size n = 7 and $\bar{X}$ a... R are computed for each sample. After 35 samples it is found that

$$\Sigma \bar{X}_i = 7805 \quad \text{and} \quad \Sigma R_i = 1200$$

(a) Setup $\bar{X}$ and R charts using these data.

(b) Assuming that both charts exhitit control, estimate the process mean and standard deviation.

(c) If the quality characteristic is normally distributed and if the specifications are 220 ± 35, can the process

Contd..... 4/-

Contd..... Q.No.4.(c)

meet the specifications ? Estimate the fraction defective.

(d) Assuming the variance remains constant, where should the process mean be located to minimise the fraction defective? What would be your estimate of fraction defective under this condition ?

((6+2)+2+(4+5)+(1+5)) = [25]

Group B

5. Suppose that a single sampling plan with n = 150 and c = 2 is being used for receiving inspection. The lot size is relatively large around 3000.

(a) Draw the OC curve of the plan by computing the probabilities of acceptance at lot qualities,

0.2%., 0.6%., 1.0%., 1.6%., 2.0%., 2.4%., 2.8% and 3.6% .

(b) Draw the AOQ curve and find the AOQL.

((5+8)+(5+4+3)) = [25]

6.(a) What is an acceptance sampling plan ?

(b) Explain the terms, (i) AQL, (ii) LTPD, (iii) α and (iv) β .

(c) Find an appropriate Dodge-Romig Single Sampling plan for lot size N = 5000, to have an AOQL of 2.5% where the rendors process average is at most 1.0%.

(d) Draw the OC curve for this plan.

(2+8+5+10) = [25]

: bcc:

1986-87	E262

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications : 1986-87

PERIODICAL EXAMINATION

PART - II

Design of Experiments (Theory and Practical)

Date: 20.5.1987 Maximum Marks: 100 Time: 2 hours

Note: Attempt all questions.

1.(a) Explain the three basic principles of design.

(b) Give layout and analysis of CRD. (15+15) = [30]

2.(a) Write down analysis of variance model for a two-way classified data with one observation percell.

(b) For the above model show that MSE provides an estimate of error sum of squares.

(c) How will you estimate confidence interval for the difference of two treatment effects ?

(5+15+10) = [30]

3. Below are given the yields in gm. per plot (plot size = $\frac{1}{2000}$ acre) for the varieties of seed cotton :

Variety 1	Variety 2	Variety 3
77	109	46
70	106	70
63	137	71
84	79	65
95		61

(a) Write out the analysis of variance table.

(b) Test if the varieties differ significantly among themselves.

(c) If the result of (b) is affirmative, determine which varities differ significantly.

(5+20+15) = [40]

1986-87 E252

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications: 1986-87

PERIODICAL EXAMINATION

Sample Surveys (Theory and Practical)

Date: 6.5.1987 Maximum Marks: 100 Time: 2 hours

Note: Answer any FOUR questions.

1.(a) Compare ratio estimator with the mean per unit estimator - under simple random sampling without replacement. Obtain an expression for bias of the ratio estimator.

(b) The following table gives the present population (x) and past population (y) for a list of 20 villages. Draw a SRSWOR of 3 villages and on the basis of the sample obtain ratio estimate of present total population. Also obtain coefficient of variation of the estimator.

Village sl. no.	Present popn.(x)	Past popn. (y)
1	25	17
2	91	78
3	13	27
4	62	57
5	71	74
6	67	65
7	55	84
8	91	73
9	103	92
10	29	25
11	72	67
12	62	57
13	71	54
14	81	64
15	41	32
16	72	78
17	109	5
18	212	200
19	73	64
20	5	82

(5+8+12) = [25]

p.t.o.

2.(a) What is an Interpenetrating Network of Sampling ? State how by using this technique you can get an unbiased estimator of the variance of an estimator for a population Hence or otherwise obtain an unbiased variance estimator in Systematic Sampling.

(b) The following table gives the no. of workers (x) value added by manufacture (y) of 21 factories situated in North Calcutta. Draw a PPSWR sample of size 5 by Lahiri's method and obtain an estimate of the averages value added by manufacture for all the 21 factories. Obtain an estimate of the variance of the estimator.

Factory sl. no.	No. of workers ('00 Rs.)(x)	Value added ('000 Rs.)(y)
1	1.5	5.7
2	.7	2.3
3	4.1	11.7
4	.9	2.9
5	8.1	21.4
6	41.2	100.3
7	6.1	15.7
8	8.2	21.8
9	7.1	19.7
10	5.9	23.2
11	6.2	15.3
12	6.1	17.2
13	5.2	12.9
14	4.2	11.8
15	3.8	10.9
16	2.1	5.3
17	8.1	21.8
18	6.3	17.2
19	7.2	19.3
20	6.5	20.1
21	3.4	11.3

(4+6+4+11) = [25]

3. Explain the concept of two stage sampling. Obtain an unbiased estimator for population mean under this scheme of sampling and derive expressions for its variance under

Contd..... 3/-

Contd..... Q.No.3

simple random sampling without replacement at both the stages.

Obtain an unbiased estimator for the variance of the estimator for population total 6 where sampling is by (a) SRSWR at first stage (b) PPSWR at second stage.

(5+5+8+3+4) = [25]

4.(a) Under the usual assumptions, which you are to state explicitly, prove that

$$V_{ran} \leq V_{prop} \leq V_{opt} .$$

Can it happen that $V_{prop} > V_{ran}$? Explain.

(b) It happened that in a particular stratum h, $n_{h(opt)} > N_h$. How would you find sample sizes under optimum allocation under such a situation. Find an unbiased estimator for population mean and an expression for its variance.

(c) When is a two stage sampling design said to be self-weighted ?

(8+5+8+4) = [25]

5.(a) In SRSWOR it is specified that the estimate of population mean should lie within an interval of length d from the true value. Obtain an expression for sample size necessary to achieve this result.

(b) Write short notes on

(i) advantages and disadvantages of a sample survey vis-a-vis complete enumeration.

(ii) use of SRSWR vis-a-vis SRSWOR in estimating population total.

(iii) use of systematic sampling procedure.

(10 + 5 x 3) = [25]

:bcc:

1986-87	E254

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications: 1986-87

PERIODICAL EXAMINATION

PART - II

Sample Surveys (Theory and Practical)

Date: 29.4.1987 Maximum Marks: 100 Time: 2 hours

Note: Answer any FOUR questions. Each question carry equal marks. Maximum Score for each subdivision of a question has been indicated in the margin.

1.(a) Explain and illustrate the concept of Simple Random Sampling Without Replacement. Find an unbiased estimator of population mean under this sampling scheme and show that an unbiased estimator of the variance of this estimator exists.

(b) A simple random sample (without replacement) of 30 scripts for mathematics of ICSE students was taken from a total of 12328 answer scripts an the subject and the percentage marks in these scripts are given below.

65, 73, 34, 95, 87, 99, 62, 75, 74,
61, 82, 25, 92, 95, 89, 36, 41, 53,
64, 75, 87, 96, 89, 71, 60, 60, 52,
34, 91, 87.

Estimate the average marks in mathematics for all the above 12328 students. Estimate the standard error of this estimator. Obtain 95% confidence interval for the average score in the population.

[(5+6+6) + (1+4+3)] = [25]

2.(a) What are the principal reasons and advantages of stratification ? How should one stratify so that these advantages are most suitably utilised ?

Contd..... 2/-

Contd..... Q.No.2

(b) Obtain the formula for Neyman's allocation of sample size in various strata assuming a linear cost function. Also obtain the value of total sample size when objective is to (i) minimise variance for a fixed cost (ii) minimise cost for a fixed variance.

(c) The following table gives the diameter of 323 lead balls. Draw a simple random sample of 9 balls.

Diameter (mm)	No. of balls
13	31
14	72
15	37
16	81
17	15
18	62
19	25

$[6+(8+3+3)+5] = [25]$

3.(a) Considering clusters of equal sizes find an unbiased estimator of population mean in clustor sampling. Find expressions for its variance in terms of population intra-class correlation coefficient among units belonging to the same cluster. Hence find the efficiency of cluster sampling estimator w.r.t. mean/unit estimator in a comparable simple random sampling (SRS) procedure. When is the cluster sampling more precise than the SRS ? Compare the schemes from the point of view of cost of data collection.

(b) The following table gives the value added by manufacture of 41 factories situated in North Calcutta and adjacent areas. Draw two independent circular systematic samples of size 7 each and find an estimator of the mean value added by all these 41 factories. Find an estimator of the variance of this estimator.

Contd..... 3/-

Contd..... Q.No.3.(b)

Factory sl. no.	Value added ('000 Rs.)	Factory sl. no.	Value added ('000 Rs.)
1	5.7	22	12.9
2	2.3	23	3.7
3	11.7	24	11.8
4	2.9	25	12.5
5	21.4	26	7.2
6	100.3	27	6.4
7	15.7	28	11.8
8	21.8	29	9.3
9	19.7	30	15.7
10	23.2	31	18.2
11	15.3	32	27.5
12	17.2	33	3.9
13	12.9	34	12.1
14	11.8	35	18.5
15	10.9	36	26.7
16	5.3	37	18.7
17	21.8	38	19.2
18	17.2	39	10.2
19	19.3	40	15.7
20	20.1	41	16.2
21	11.3		

[(2+6+5+2+2)+(3+2+3)] = [25]

4.(a) The following table gives the stratification of all farms in a country by farm size and the acreages of corn (maize) in the sampled farms. Obtain the estimate of average acres of corn per farm in the population and its variance Also obtain the variances of the estimates of population mean under (i) proportional allocation,(ii) optimum allocation for the same total sample size of 27 farms. Hence obtain the efficiency of the given allocation and the proportional allocation with respect to optimum allocation.

Contd..... 4/-

Contd..... Q.No.4.(a)

Farm size (acres)	Number of farms N_h	Acreages of corn in sampled farms ('0 acres)
0 - 40	12	3.5, 2.7, 1.1, 1.9
41 - 80	24	2.4, 6.9, 3.7, 2.8, 3.4, 2.6
81 - 120	19	1.0, 1.3, 2.9, 1.6, 3.4
121 - 160	13	1.0, 0.0, 2.3, 0.5
161 - 200	9	5.8, 4.9

Standard deviations ('0 acres) (From a past survey)
5.2
3.7
6.5
3.4
2.7

(b) A simple random sample of size 3 is drawn from a population of size N with replacement. Show that the probabilities that the sample contains 1, 2 and 3 different units (for example, aaa, aab, abc, respectively) are

$$P_1 = \frac{1}{N^2}, \quad P_2 = \frac{3(N-1)}{N^2}, \quad P_3 = \frac{(N-1)(N-2)}{N^2}.$$

[(3+5+5+6+2)+4] = [25]

5.(a) Explain and illustrate the concept of probability proportional to size sampling with replacement. Find an unbiased estimator of population mean under this sampling scheme. Find an unbiased estimator of the variance of this estimator.

(b) Obtain an estimator of population mean which utilises the values of an auxiliary variable closely related to the main variable under study. Derive large sample expressions of its variance.

[(5+5+5)+(5+5)] = [25]

:bcc:

INDIAN STATISTICAL INSTITUTE | 1986-87 | E242

One Year Evening Course in Statistical Methods and Applications: 1986-87

PERIODICAL EXAMINATION

PART - II

Statistical Inference (Theory and Practical)

Date: 22.4.1987 Maximum Marks: 100 Time: 2 hours

Note: You may answer any part of any question.

1.(a) Describe briefly the paired t-test and Fisher's t-test ?

(10+10) = [20]

(b) The heart weights in grams of 12 female and 12 male cats are given below. Does the heart of a male cat on an average weigh more than that of a female cat ?

Males : 12.7, 15.6, 9.1, 12.8, 8.3, 11.2, 9.4, 8.0, 14.9, 10.7, 13.6, 9.6.

Females : 7.4, 7.3, 7.1, 9.0, 7.6, 9.5, 10.1, 10.2, 10.1, 9.5, 8.7, 7.2.

[15]

2.(a) Describe clearly the testing problem of a one-way classified data; write down the testing procedure as well as the ANOVA table for the same problem.

(10+15) = [25]

(b) Evaluate the expected values of MSB and MSW in a one-way classified data.

(7+10) = [17]

3. The following table gives the weights in grams of the anterior muscles of both hind legs of 16 normal rabbits:

p.t.o.

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications: 1986-87

PART - I'

SUPPLEMENTARY EXAMINATION

Descriptive Statistics

Date: 8.4.1987 Maximum Marks: 100 Time: 2 hrs.

Note: Answer as many questions as you can, but the maximum marks allotted will be 100.

1.(a) Show that cumulative frequency graphs intersect at median.

(b) Define central tendency and dispersion concept for a set of data. Also, show that mean deviation $\frac{1}{n}\Sigma|x_i-A|$ is minimum when A is equal to median.

(4+4+8) = [16]

2.(a) For two sets of data with n_1 and n_2 number of observations, if means are $\bar{x}_1$, $\bar{x}_2$, standard deviations are s_1 and s_2, so that the combined variance is given by

$$s^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{(n_1 + n_2)} + \frac{n_1 d_1^2 + n_2 d_2^2}{(n_1 + n_2)}$$

where $d_1 = \bar{x}_1 - \bar{x}$, $d_2 = \bar{x}_2 - \bar{x}$, $\bar{x}$ being the combined mean.

(b) Show that, if the variable takes the values 0, 1, 2,.. with frequencies proportional to the binomial coefficien $n_{c_1}, n_{c_2}, \ldots, n_{c_n}$ respectively, then mean of the dist bution is $\frac{n}{2}$, and mean square deviation about $x = 0$ i $\frac{1}{4}.n(n+1)$.

(8+8) = [16]

3.(a) Define binomial distribution and find its mode.

(b) Show that binomial distribution with parameters n and p tends to poisson distribution provided $n \to \infty$ and $p \to 0$ such that np is finite.

Contd..... 2/-

Contd..... Q.No.3

(c) A perfect die is thrown a large number of times in sets of 8. The occurance of 5 and 6 is called a success. In what proportion of the sets would you expect to get 3 successes.

(2+4+8+2) = [16]

4.(a) Define standard normal deviate.
If $X \sim N(0,1)$, find expectation and variance of $Y = X^2$

(b) Steel rods are manufactured to be 3 inches in diameter, but they are acceptable if they are inside the limits 2.99 inches and 3.01 inches. It is observed that about 5% are rejected for over size and 5% are rejected for under size. Assuming that diameters are normally distributed, calculate what proportion of rejects would be if the permmssible limits were widened to 2.985 inches and 3.015 inches.

(3+6+7) = [16]

5.(a) Write a short note on the role of scatter diagram for bivariate distributions.

(b) The variates X and Y are correlated with correlation coefficient ρ. Two variates $U = X \cos\alpha + Y \sin\alpha$ and $V = Y \cos\alpha - X \sin\alpha$ will be uncorrelated if

$$\tan 2\alpha = \frac{2\rho\sigma_x\sigma_y}{\sigma_x^2 - \sigma_y^2},$$

where σ_x^2, σ_y^2 are variances of x and y.

(c) Show that for a set of bivariate data, the arithmatic mean of two regression coefficients is greater than the correlation coefficient between them.

(6+7+3) = [16]

6. 12 dice were thrown 2.630 times and each time the number of dice which had 5 or 6 on the upper most faces was recorded. The results are given in the following table:

Contd..... 3/-

Contd..... Q.No.6

No. of dice with 5 or 6 upper most	0	1	2	3	4	5	6	7	8	9	10	11	12
Frequency	18	115	326	548	611	519	307	133	40	11	2	0	0

Graduate the observed distribution with a binomial distribution for which p is unknown.

[20]

7. Find correlation coefficient between two variables x and y having values, and state whether x and y are independent.

X	-5	-4	-3	-2	-1	0	1	2	3	4	5
Y	25	16	9	4	1	0	1	4	9	16	25

[20]

:bcc:

1986-87 | E241

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications : 1986-87

PERIODICAL EXAMINATION

PART - II

Inference

Date: 1.4.1987 Maximum Marks: 100 Time: 2 hrs.

Note: You may answer any part of any question.

1.(a) Define an unbiased estimator and minimum variance unbiased estimator of a parametric function.

(3+5) = [8]

(b) What is meant by saying that a minimum variance unbiased estimator, if exists, is unique ? Prove your statement.

(5+15) = [20]

(b) If Y has the Binomial distribution $B(n ; \theta)$ where n is a known positive integer and θ can take any value in $[0 , 1]$, of what quantity is Y^2 an unbiased estimator ? Find also an unbiased estimator of the variance of Y.

(10+10) = [20]

2. If $T_1, \ldots, T_k$ are unbiased estimators of θ having the same positive variance and θ can take any real value, find the linear combination of the form

$$S = a + b_1 T_1 + \ldots + b_k T_k$$

such that $a_1 b_1, \ldots, b_k$ are real constants, S is unbiased for θ and S has the smallest variance among all such linear combinations.

[20]

3.(a) Define the power, size and level of significance of a test procedure.

[9]

Contd..... 2/-

Contd..... Q.No.3

(b) Write down clearly the critical region of the standard test procedure for testing $H_0 : \mu = \mu_0$ against $H_1 : \mu > \mu_0$ based on a sample from a normal population with the unknown mean μ and the known variance $\sigma^2 > 0$. [7]

(c) How does the critical region in (b) change in case the alternative hypothesis becomes $H_1 : \mu < \mu_0$. [3]

(d) Draw the graph of the power function of the test procedure in (b) when $n = 500$ and $\sigma^2 = 1$. [30]

(e) Draw the graph of the power function of the test procedure in (c) when $n = 500$ and $\sigma^2 = 1$. [6]

(In (d) and (e), take $\alpha = .01$)

:bcc:

1986-87	E134

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications: 1986-87

FINAL EXAMINATION

PART - I

Economics / Statistics : (c) Time Series
(d) Official Statistics

Date: 12.1.1987 Maximum Marks: 100 Time: 2 hrs.

Note: Answer any FIVE questions. All questions carry equalmarks.

1. The following table gives the data on the wholesale price index numbers of cereals and manufactures (with base 1970-71 = 100) for each of the years 1971-84:

Year	Price Index of cereals (Rs.)	Price Index of manufactures
1971	100.0	100.0
1972	111.4	119.0
1973	128.8	133.8
1974	178.1	163.1
1975	185.7	173.1
1976	151.6	171.7
1977	161.2	179.4
1978	158.1	177.9
1979	167.0	203.9
1980	189.5	249.9
1981	213.1	269.5
1982	229.4	269.6
1983	259.2	287.9
1984	246.9	313.3

The relative prices of cereals is defined to be the ratio of P_c to P_m.

(a) Draw a diagram showing the movement in the relative price of cereals over the years 1971-84.

(b) Which components of a time series are present in the diagram in (a)? comment on the behaviour of these components.

p.t.o.

2. Obtain the deseasonalised series from the following series by obtaining seasonal indices through 'link relative method' or 'ratio to moving average' method :

Year	Quarters			
	I	II	III	IV
1962	1.76	1.33	1.48	1.71
1963	2.04	1.39	1.40	1.49
1964	2.16	1.43	1.58	2.05
1965	2.03	2.10	1.79	2.52

3. Consider the following residual series obtained from a 3-year moving average applied to agricultural income of Indian Union. Draw a graph for the series and test whether the series is purely random. Do you think that the sample series is sufficiently large to permit you to obtain a firm inference ?

t	1901-02	02-03	03-04	04-05	05-06
Residual	- 29.0	25.0	10.4	-26.3	-0.3
t	06-07	07-08	08-09	09-10	10-11
Residual	45.0	-57.0	1.7	24.4	3.4
t	11-12	12-13	13-14	14.15	15.16
Residual	-6.3	-1.7	-22.3	-3.3	7.4
t	16-17	17-18	18-19		
Residual	14.7	30.0	-69.6		

4. Fit a linear and quadratic trend to the following data and determine which of the two gives a better prediction, use some standard yardstick to measure the in accuracy in prediction.

Contd..... 3/-

Contd..... Q.No.4

Year	Income	Year	Income
1950-51	85	1956-57	110
51-52	91	57-58	109
52-53	95	58-59	117
53-54	100	59-60	129
54-55	103	60-61	127
55-56	105	61-62	130

While comparing the accuracy in prediction by the two models, you may consider the results for the years 1962-63 to 1965-66 only.

5. Fit an exponential growth curve to the following series of agricultural production Index of India:

Time	Ag. Production Index
1	142.2
2	144.8
3	139.6
4	143.1
5	159.4
6	133.1
7	131.6
8	161.1
9	159.5
10	170.8
11	182.2

6. Write short notes on any one of the following:

(i) Agriculture Statistical System in India.

(ii) Government Statistical Organisations (at the Centre) in India.

:bcc:

1986-87 | E133

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications: 1986-87

FINAL EXAMINATION

PART - I

Economics Statistics: (a) Index Number
(b) National Income Accounting

Date: 9.1.1987 Maximum Marks: 100 Time: 2 hours

Note: Answer any FIVE questions.

1. The gross domestic product of Indian economy at constant 1960-61 prices and at current prices are given below.

Year	G.D.P. ($Rs.10^9$) (at constant 1960-61 prices)	G.D.P. (at current prices)
1960-61	70.10	70.10
1961-62	70.88	72.57
1962-63	69.10	74.13
1963-64	71.16	82.91
1964-65	77.49	104.71
1965-66	66.60	101.28

Compute the appropriate price and quantity indices with 1964-65 as the base.

[16]

2. Combine the two series of Index numbers given below to obtain a new series with (a) 1963 as base (b) 1960 as base

Year	Old series 1958 = 100	New series 1963 = 100
1960	101	
1961	105	
1962	112	
1963	126	100
1964	134	106
1965		99
1966		102

[16]

p.t.o.

3. The following table gives the change in the price and the consumption (quantity) of certain major constituents of the consumption basket of the labor class:

Commodity (unit)	1950		1960	
	Price	Consumption (quantity)	Price	Consumption (quantity)
1) Wheat (Quintals)	100	1000	110	1000
2) Rice (Quintals)	150	15	170	18
3) Cloth (Metrics)	5	50	4	30

(a) What should be the increase in wage rate in order to maintain the standard of living of the labour class at the 1950 level ?

(b) What should have been the wage rate in 1950 if the labour class were to maintain the level of living of 1960 in 1950 ?

[16]

4. In the economy described below what is GNP (a) by category of final demand at market prices

(b) by industry of origin at factor cost.

(a) A sells to B for Rs.50.00 and to C for Rs.50.00. B sells to private consumption for Rs.40.00 and to export for Rs.80.00. C sells to capital formation for Rs.50.00.

(b) A sells to private consumption for Rs.60.00. B sells to C for Rs.30.00 and to export for Rs.50.00; C, who also imports for Rs.30.00, sells to capital formation for Rs.70.00.

(c) A imports for Rs.50.00 and sells to exports for Rs.20.00 and to B for Rs.40.00. B sells to private consumption for Rs.60.00. C sells to D for Rs.40.00 and to capital formation for Rs.10.00. D sells to exports for Rs.50.00 and to public consumption for Rs.20.00.

(d) Same as (B), but an indirect tax of 20% is imposed on A's sales, an indirect tax of 50% on B's sales to C and a custom duty of $33\frac{1}{3}$% is imposed on imports. All these taxes and duties are 'passed on' in price increase on the goods concerned.

contd..... 3/

Contd..... Q.No.4

(e) A sells to B for Rs.20.00, B sells to private consumption for Rs.40.00; C buys imports for Rs.30.00 and sells to private consumption for Rs.60.00. A subsidy of $33\frac{1}{2}\%$ is granted on C's sales to private consumption.

[26]

5. Write short notes on any three of the following:

(a) Chain index (b) True cost of living Index
(c) Social accounting Matrix (d) Make matrix and the construction of commodity r commodity Input-output table.

[26]

6. The table below gives the retail prices of the group 'food grains and products' in the two centres; Calcutta and Bombay in 1960. Column 2 of this table gives the average family expenditure (of industrial workers) in Calcutta in 1960. Using these informations, calculate the consumer price index number of 'food grains and products for Bombay taking Calcutta as base:

Item	Average family exp. in Calcutta in 1960 (Rs.)	Unit	Price (Rs.) in	
			Calcutta	Bombay
(1)	(2)	(3)	(4)	(5)
Rice	17.55	Kg	0.74	0.70
Wheat	4.21	Kg	0.40	0.41
Bread	0.21	500 Gm	0.40	0.48
Arhar Dal	0.92	Kg	0.75	0.76
Gram Dal	0.27	Kg	0.58	0.60
Moog Dal	0.77	Kg	0.78	0.90
Musur Dal	0.79	Kg	0.75	0.78

[15]

1986-87 | E124

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications : 1986-87

FINAL EXAMINATION

PART - I

Descriptive Statistics (Practical)

Date: 7.1.1987 Maximum Marks: 100 Time: $2\frac{1}{2}$ hrs.

Note: Answer all questions. Marks allotted to each question are given in brackets [ ].

1. The time required by a nurse to inject a shot of penicillin has been observed to be normally distributed with mean $\mu = 30$ seconds and a variance $\sigma^2 = 100$ square seconds. Find the following percentiles:

(a) 10th (b) 50th and (c) 75th

Hence, find (d) 25th and (e) 90th percentiles.

[8]

2. The following constants are obtained from measurements on length in mm.(x_1), volume in c.c.(x_2) and weight in gm.(x_0) of 300 eggs:

$\bar{x}_0 = 56.03$ $\bar{x}_1 = 55.95$ $\bar{x}_2 = 51.48$

$s_0 = 4.41$ $s_1 = 2.26$ $s_2 = 4.39$

$r_{01} = 0.581$ $r_{02} = 0.974$ $r_{12} = 0.578$

(i) Obtain the linear regression equation of egg-weight on egg-length and egg-volume. Hence, estimate the weight of an egg whose length is 58.0 mm. and volume is 52.5 c.c.

(ii) Calculate the multiple correlation coefficient assuming that x_1 and x_2 are independent variables. Also, find the coefficient of determination and comment on it.

(22+5) = [27]

p.t.o.

3. During an investigation in an agricultural farm in Bengal, the length (in cm.) of green jute plant (x) and the weight (in gm.) of dry jute fibre (y) were observed for 350 plants. With these data the following bivariate frequency table was obtained:

y \ x	111.5	127.5	143.5	159.5	175.5	191.5	207.5
1.175	12	25	15	1	-	-	-
2.775	1	4	33	59	29	3	-
4.375	1	-	4	28	35	14	2
5.975	-	-	-	2	20	18	1
7.575	-	-	-	1	1	14	5
9.175	-	-	-	-	4	8	2
10.775	-	-	-	-	-	3	2
12.375	-	-	-	-	-	-	3

(i) From the above table, compute the coefficient of correlation of the two variables. Also, find the linear regression equation of weight of dry fibre on length of green plant.

(ii) Determine the correlation ratio of y on x and comment on it.

(50+15) = [65]

:bcc:

1986-87 | E123

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications : 1986-87

FINAL EXAMINATION

PART - I

Descriptive Statistics (Theory)

Date: 5.1.1987 Maximum Marks: 100 Time: 2 hrs.

Note: Attempt as many questions as you can, but the maximum marks allotted would be 100.

1.(a) Define with illustration the following:

(i) Histogram

(ii) Frequency polygon and Frequency curve

(iii) Cumulative frequency curve. [10]

(b) A population of values is symmetrically distributed about the constant k. Examine whether the following statements are correct:

(i) the mean is k

(ii) the mean coincides with the mode

(iii) the distribution can not be binomial

(iv) $\frac{1}{2}$ (upper quartile + lower quartile) is equal to k

(v) the median coincides with mean. [15]

2.(a) Show that the sum of the absolute deviations from the median is a minimum. [10]

(b) Show that if $\bar{x}_1$ and $\bar{x}_2$ are the means of two sets, then the mean $\bar{x}$ of the combined set also lies between $\bar{x}_1$ and $\bar{x}_2$. [5]

(c) In a frequency table, the upper boundary of each interval has a constant ratio to the lower boundary. Show that the G.M. (G) may be expressed by the formula

$$\log G = X_o + \frac{C}{N} \Sigma f_i (i-1)$$

where X_o is the logarithm of the mid value of the first interval and C is the logarithm of the ratio between the upper and lower boundaries. [10]

p.t.o.

3.(a) Find the most probable number of successes in a series of n independent trials, the borbability of each success being p. [5]

(b) Find the mode of the poisson distribution. [5]

(c) If x is a poisson variate with mean m, what would be the expectation of e^{-kx}, kx, where k is a constant ? [5]

4.(a) What is the probability density function of the Binomial distribution ? What form does it take, when, n, the number of trials, is very large and neither p nor q is very small ? [10]

(b) What is the mean deviation from the mean m of the normal distribution with parameters m and σ ? Show that all odd order moments about the mean m also vanish. [10]

5.(a) Define product moment Correlation Coefficient. Does it remain unchanged, if the old variables are transformed to new variables by change of origin and scale.

Show that, Correlation Coefficient r, satisfies

$$|r| \leq 1$$

always. [10]

(b) Show that θ, the acute angle between the two lines of regression, is given by

$$\tan\theta = \frac{1 - r^2}{r} \cdot \frac{\sigma_x \sigma_y}{\sigma_x^2 + \sigma_y^2}$$

where, σ_x and σ_y are the standard deviations of x and y and r is the Correlation Coefficient between x and y. [10]

6.(a) Define

(1) Multiple Correlation Coefficient and

(2) Partial Correlation Coefficient.

(b) Suppose we are given N sets of corresponding values of three variates x_1, x_2 and x_3. Let these variates be measured from their respective means and the quantities so obtained be denoted by x_1, x_2 and x_3, what should be the regression equation of x_1 on x_2 and x_3 ? [10]

(c) Prove the identity

1986-87	E112

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Application : 1986 - 87

PERIODICAL EXAMINATION

Probability .

Date: 26.11.1986 Maximum Marks: 100 Time: 2 hrs.

Note: The paper carries 120 marks. Answer as much as you can.

1.(a) Define density function of a discrete random variable and state its properties.

(4+4) = [8]

(b) Suppose X is a discrete random variable having density and given by

x	-3	-1	0	1	2	3	5	8
f(x)	.1	.2	.15	.2	.1	.15	.05	.05

Find (i) P (X is negative)

(ii) P ($1 \leq X \leq 8$)

(iii) P ($X = -3 | X \leq 0$)

(iv) P ($X \geq 3 | X > 0$)

(3+3+3+3) = [12]

2.(a) Define the distribution function of a discrete random variable. Find the density f of a random variable X following the geometric distribution with parameter p, and also find the distribution function of X.

(3+7) = [10]

(b) A die is rolled until a 6 appears.

(i) What is the probability that at most six rolls are needed ?

(ii) How many rolls are required so that the probability of getting 6 is at least 1/2 ?

(5+5) = [10]

(c) Let X and Y be independent random variables each geometrically distributed with parameter p and $P(X \geq x) = (1-p)^x$ for the nonnegative integer x.

Contd..... 2/-

Contd..... Q.2.(c)

(i) Find $P(Y \geq X)$

(ii) Find the distribution of X+Y.

(5+5) = [10]

3.(a) Let X and Y be two discrete random variables, having the joint density f. Find the marginal densities of X and Y [10]

(b) Suppose X and Y are random variables that assume the values x and y, with probabilities given by the table:

Y \ X	0	1	2	3
1	$\frac{1}{8}$	0	0	$\frac{1}{8}$
2	0	$\frac{2}{8}$	$\frac{2}{8}$	0
3	0	$\frac{1}{8}$	$\frac{1}{8}$	0

Find the marginal densities of X and Y ; and also determine wheather the random variables are independent.

(10+5) = [15]

4. Find the distribution of the sum of r independent, identically distributed random variables each following the geometric distribution with parameter p. [20]

5.(a) Define the expectation of a discrete random variable. [5]

(b) Let X and Y be two random variables having finite expectation.

(i) Show that X+Y has finite expectation and E(X+Y) = EX+EY.

(ii) Suppose $P(X \geq Y) = 1$. Then show that $EX \geq EY$; moreover EX = EY if and only if $P(X = Y) = 1$.

(10+10) = [20]

:bcc:

INDIAN STATISTICAL INSTITUTE | 1986-87 | E132

One Year Evening Course in Statistical Methods and Applications: 1986-87

PERIODICAL EXAMINATION

Economic Statistics (Theory and Practical)(c+d)

Date: 3.12.1986 Maximum Marks: 100 Time: 2 hrs.

Note: Marks for each part of each question are indicated in the brackets [ ]. You may attempt any part of any question.

1. Write short notes:

   (a) Agricultural Statistical System in India with particular reference to land utilisation and yield statistics. [12½]

   (b) National Sample Survey Organisation. [12½]

2. (a) The values of cost of living indices in country A for various years between 1945 and 1952 with 1948 = 100 were as follows:

   93 93 96 100 102 103 119 and 128

   Obtain the changed values of these indices by shifting the base to 1951. [12½]

   (b) Discuss 'Factor reversal test' and 'Time reversal test'. [12½]

3. The table below relates to the weekly pay of the wage camers on a company's pay roll:

Category of employees	1950 Number	1950 Total(Rs.) pay	1960 Number	1960 Total(R. pay
Officers	350	5000	300	5500
Skilled workers	1400	7000	1500	10000
Unskilled workers	3000	10000	2500	13000

Construct an index of weekly earnings for 1960 based on 1950. [18¾]

p.t.o.

4. Transactions in an economy are basically as follows:

Industry A sells to Industry B for £ 50 ;

Industry B sells to Industry C for £ 80 ;

Industry C sells to final demand for £ 120.

(i) What is total national product at market prices and national product at factor cost, if the transactions are modified as follows (each are separately) and assuming that all indirect taxes or subsidies are 'passed on' fully as price increases throughout the economy. [6¼]

(ii) as for (i) but an indirect tax of one-third is also imposed on the product of industry C (i.e., on the value of his sales including his allowance for the higher price he had to pay as a result of (i).

[6¼]

(iii) an indirect tax of 20% is imposed on the products of industry A which adds the tax to its price (to industry B). But Industry B cannot pass on the price increase and absorbs the whole Tax.

[6¼]

(iv) a subsidy of £ 20 is given to Industry B which enables it to reduce the value of its sales to Industry C from £ 80 to £ 60, and Industry C reduces the value of its sales to private consumption from £ 120 to £ 100.

[6¼]

5. Initially the following, and only the following transactions take place in the economy :

A sells to B for £ 50
B sells to C for £ 80
C sells to private consumption.

(i) What is GNP by industry of origin ? [6¼]

(ii) What is GNP by industry of origin if now, A imports for £ 20 ? [6¼]

Contd..... 3/-

Contd..... Q.No.5

(iii) What is GNP by industry of origin if A imports for £ 10, B imports for £ 10 ? [$6\frac{1}{4}$]

(iv) What is GNP by industry of origin, and by category of final demand if A imports for £ 20 and C sells, in addition to his sales to private consumption indicated above, £ 30 to exports ? [$6\frac{1}{4}$]

:bcc:

1986-87 | E122

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications: 1986-87

PERIODICAL EXAMINATION

Descriptive Statistics (Theory and Practical)

Date: 19.11.1986 Maximum Marks: 100 Time: 2 hrs.

Note: Attempt all the questions.

1.(a) Deduce the first four moments about the mean of the normal distribution. [5]

(b) If X is normally distributed with mean 3 and variance 1, find the expectation and variance of e^{aX}. [5]

(c) The life (in hours) of electronic tubes of a certain type is supposed to be normally distributed with mean $\mu = 155$ hour and s.d. = 19 hour. What is the probability that the life of a tube will be

(i) between 136 hours and 174 hours ?

(ii) between 117 hours and 193 hours ?

(iii) less than 119 hours ?

(iv) more than 190 hours ? [10]

(d) Fit a normal distribution from the following frequency distribution of weights in pounds of 200 babies at birth:

Weights (in pounds)	Frequency
2.1 - 2.7	1
2.7 - 3.3	6
3.3 - 3.9	18
3.9 - 4.5	35
4.5 - 5.1	46
5.1 - 5.7	45
5.7 - 6.3	28
6.3 - 6.9	15
6.9 - 7.5	5
7.5 - 8.1	1
	200

[40]

p.t.o.

2.(a) Define Correlation Coefficient between two variables and give one example that two uncorrelated variables may not be independent.

Deduce that $-1 \leq r \leq +1$ [10]

(b) x_1 and x_2 are two variates with variances σ_1^2 and σ_2^2 respectively and r is the correlation between them. Determine the value of k such that $(x_1 + kx_2)$ and $x_1(\sigma_1/\sigma_2)\ x_2$ are uncorrelated. [10]

3.(a) Find the most likely price in Bombay (x) corresponding to the price of Rs.70/- at Calcutta (y) from the following data :

Average price at Calcutta	= 65
Average price at Bombay	= 67
S.D. at Calcutta	= 2.5
S.D. at Bombay	= 3.5

Coefficient of Correlation is 0.8 between the two prices of the commodity in two towns. [5]

(b) In a partially destroyed Laboratory record of an analysis of correlation data, the following results only are legible :

Variance of x = 9

Regression Equations : $8x - 10y + 66 = 0$;
$40x - 18y = 214$.

What were

(a) the mean values of x and y
(b) the standard deviation of y

and (c) the coefficient of correlation between x and y. [15]

:bcc:

1985-86	E253S

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications : 1985-86

SUPPLEMENTARY EXAMINATION

Sample Surveys (Theory and Practical)

Date: 5.11.1986 Maximum Marks: 100 Time: 2 hrs.

Note: Attempt as many questions as you can. The paper carries 110 marks but the maximum score is to be 100 marks. Marks are specified at the end of each question.

1. (i) Define ratio, product, difference and regression estimators, for population mean $\bar{Y}$ in case $\bar{X}$ (the mean of an auxiliary character x) is known.

(ii) Mention the situations under which above estimators, based on SRSWOR, are to be preferred over the sample mean.

(iii) A simple random sample of n = 25 households from a village consisting of 300 households was selected and size x of the selected household and its annual earning (y) observed. The data yielded

$$\Sigma x_i = 160 \; ; \quad \Sigma y_i = 155X.10^3 \; ; \quad \Sigma x_i^2 = 1625$$

$$\Sigma y_i^2 = 135 \times 10^7 \; , \quad \Sigma x_i y_i = 1232 \times 10^3$$

It is known that average household size for the village is 8.

(a) Obtain ratio, regression and product estimates for the average annual earnings per household in the village.

(b) Obtain estimated percent relative efficiencies of the above estimates over the sample mean.

(c) Estimate the annual earning per individual in the village.

(10+10+15) = [35]

p.t.o.

2. (i) Define simple stratified sampling and proportional and Neyman allocations. Give estimators for population mean, total and proportion.

(ii) Define PPS sampling. Give estimators for population mean and total.

(iii) Define linear-systematic and circular systematic sampling and give unbiased estimators for population mean and total.

(iv) Define cluster sampling and give estimators for population total and mean per cluster and per element.

(10+5+10+5) = [30]

3. The member of workers in ten Iron and Steel factories are 70, 30, 25, 100, 80, 60, 150, 200, 95 and 125 respectively.

(a) Draw samples of two factories with probabilities proportional to member of workers using cumulative size method and Lahiri-method.

(b) In case of Lahiri-method, let the selected factories have their last year's profits as Rs.3 lakhs and 2 lakhs respectively. Find the average profit per factory for the ten factories under investigation.

(c) Obtain estimated standard error of the estimate.

(10+5+10) = [25]

4. A village is divided into 10 parts each consisting of 30 households. A sample of 2 parts is selected using SRSWOR and then samples of 5 households are selected, from each of the selected parts using SRSWOR. Information on number of characteristics is collected from the selected households. The following data pertain to the number of milk-yielding cattles for the selected households.

Sampled parts	No. of milk yielding cattles for selected households				
1	2	10	3	1	8
2	15	5	2	1	7

(a) Estimate total number of milk yielding cattles and number of cattles per household in the village.

(b) Estimate the standard error of above estimates.

(10+10) = [20]

1986-87 | E131

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications: 1986-87

PERIODICAL EXAMINATION

Economics Statistics (Theory and Practical)(a+b)

Date: 5.11.1986 Maximum Marks: 75 Time: 2 hrs.

Note: The paper is set for 90 marks. You may attempt any part of any question. The marks for each question are given in brackets.

1. Using the method of link-relatives, obtain the seasonal indices from the following table, assuming the trend to be linear:

Total production of paper ('000 tons)

Year \ Quarter	1st	2nd	3rd	4th
1951	37	38	37	40
1952	41	34	25	31
1953	35	37	35	41

[20]

2.(a) Discuss briefly the special characteristics of the 'Gompertz Curve' and 'Logistic Curve' in the context of time series analysis.

[10]

(b) Fit a Gompertz Curve to the following production data by the method of partial totals:

Year	Production
1951	310
1952	332
1953	326
1954	365
1955	395
1956	415
1957	431
1958	511
1959	661

[15]

p.t.o.

3. Write short notes on the following:

   (a) Autcorrelation
   (b) Testing of randomness of a time series
   (c) Moving average method

(9+8+8) = [25]

4. The quarterly data on the wholesale price index with 1958 = 100, is given below. Obtain the seasonal indices by the ratio to trend method and compute the deseasonalised series:

Quarter	Year			
	1962	1963	1964	1965
I	111	114	124	141
II	114	119	129	140
III	118	123	140	141
IV	116	122	141	149

[20]

:bcc:

1986-87 | E111

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications: 1986-87

PERIODICAL EXAMINATION

Probability

Date: 29.10.1986 Maximum Marks: 100 Time: 2 hrs

Note: Answer as many as you can. The maximum possible score is 100.

1.(a) Define the conditional probability of an event B given another event A. [5]

(b) Suppose that 40% of the population of a certain city is male. Suppose also that 50% of the male and 30% of the female smoke. Find the probability that a smoker is male. [10]

2.(a) State and prove the Baye's theorem. [15]

(b) There are three urns U_1, U_2, U_3. Urn U_1 contains one black and four white balls. U_2 contains three white and two black balls. U_3 contains one white and four black balls. One urn is selected at random. Assume that probabilities are equal that any of the three urns will be selected. A ball is drawn from an urn selected at random and observed to be black, what is the probability that the selected urn is U_3 ? [20]

3.(a) If A_1, A_2, A_3 be events, then compute the probability of the event $A = \bigcup_{i=1}^{3} A_i$; state the corresponding results for n events. [10]

(b) If A_1, A_2, ..., A_n are events (not disjoint) on some probability space, then show (using induction or otherwise) that

$$P\left(\bigcup_{i=1}^{n} A_i\right) \leq \sum_{i=1}^{n} P(A_i).$$

[10]

p.t.o.

4. Suppose n balls are distributed into n boxes so that all of the n^n possible arrangements are equally likely. Compute the probability that only box 1 is empty.

[20]

5. The mathematics departments consists of 25 professors, 15 associate professors and 35 Lecturers. A committee of 6 is selected at random from the department. Find the probability that all the members of the committee are lecturers.

[10]

6. Consider the random distribution of n balls into r boxes and find the probability that exactly k boxes are empty.

[20]

:bcc:

1985-86 | E242S

INDIAN STATISTICAL INSTITUTE

One Year Evening Course in Statistical Methods and Applications: 1985-86

SUPPLEMENTARY EXAMINATION

Statistical Inference (Theory and Practical)

Date: 29.10.1986 Maximum Marks: 100 Time: 2 hrs.

Note: Solve as much as you can. The maximum possible score is 100.

1.(a) Let $X_1, X_2, \ldots, X_n$ be a random sample from the pdf $f(x, \theta)$

Define i) An estimator for θ

ii) An asymptotically unbiased estimator for θ

iii) The maximum likelihood estimator (MLE) for θ.

iv) Cramer-Rao lower bound (CRLB) for estimating θ.

v) A consistent estimator for θ.

(b) Let $X_1, X_2, \ldots, X_n$ be a random sample from $f(x, \lambda) = \frac{1}{\lambda} e^{-x/\lambda}, x > 0$. Find the CRLB for the variance of an unbiased estimator for λ. Is $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ efficient for λ.

(c) Is the estimator $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ consistent for the parameter of a Poisson distribution.

(d) Prove that the largest order statistic $X_{(n)}$ is a sufficient statistic for the parameter of a uniform distribution, $f(x, \theta) = \frac{1}{\theta}, 0 < x < \theta$.

(e) Let the random variable X have a uniform density given by

$$f(x, \theta) = I_{[\theta-\frac{1}{2}, \theta+\frac{1}{2}]}(x), \quad -\infty < \theta < \infty.$$

Obtain the MLE based on a sample of size n.

((2+2+4+5+3)+8+8+8+8) = [48]

2.(a) Let $X_1, X_2, \ldots, X_n$ be a random sample from $f(x, \theta)$, $\theta \in \Omega$. Suppose we are interested in testing the hypothesis

$H_0: \theta \in \Omega_0$ against $H_1: \theta \in \Omega - \Omega_0$

Contd..... 2/-

Contd..... Q.No.2.(a)

Define a nonrandomised test, its size and power.

(b) State and prove Neyman-Pearson lemma.

(c) Explain the generalised likelihood ratio test (GLRT) principle.

(d) Let $X_1, X_2, \ldots, X_n$ be a random sample from $N(\mu, \sigma^2)$. Obtain the GLRT to test $H_o: \sigma^2 = \sigma_o^2$ versus $H_1: \sigma^2 \neq \sigma_o^2$, μ being unknown. Set α, the size, equal to 0.05.

(6+10+4+10) = [30]

3. In the eighth century B.C. the Etruscan Civilization was the most advanced in all of Italy. It was originally located between the rivers Arno and Tiber (the region now known as Tuscany) and spread quickly and eventually overran much of Italy. But as quickly as it came, it faded.

No chronicles of Etruscan empire have ever been found and to this day its origins remain shrouded in mystery. Research has shown that maximum head breadth of modern Italian males averages 132.4 mm. Listed in the table below are the maximum head breadths recorded for 12 male Etruscan skulls.

Maximum head breadths (mm) of 12 Etruscan males

141	146	144	141
141	136	137	149
141	142	142	147

We would like to know whether true average head breadth of an Etruscan male is equal to 132.4 or not.

Formulate it as a hypothesis testing problem and solve it for $\alpha = 0.05$ and $\alpha = 0.01$

[20]

4. Dental structure provides and effective criterion for classifying certain fossils. Not long ago a baboon skull of unknown origin was discovered in a cave in Angola. Its third molar length was 9.0 mm. Speculation arose that the baboon in question might be a missing link and belong to the genus Papio. Members of this genus have third molars that measure, on the average 8.18 mm long with a standard deviation of 0.47 mm. What would your inference be ? Set $\alpha = 0.05$.

[10]

bcc:

1986-87 | E121

INDIAN STATISTICAL INSTITUTE
One Year Evening Course in Statistical Methods and Applications : 1986-87

PERIODICAL EXAMINATION

Descriptive Statistics (Theory and Practical)

Date: 22.10.1986 Maximum Marks: 100 Time: 2 hours

Note: Attempt all the questions.

1.(a) While grouping a data what information is ignored ? What are the principles governing the choice of (a) the number of class-intervals, (b) the length of the class intervals and (c) mid points of the class intervals. [6]

(b) What is meant by Central tendency ? What are the various measures of location of a distribution and to what purposes they are used ? [6]

(c) Let there be k sets of values of x, the number of values in the k sets being $n_1, n_2, \ldots, n_k$ and the means being $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_k$ respectively. Show that the grand mean of x is

$$\bar{x} = \Sigma n_i \bar{x}_i / \Sigma n_i \qquad [5]$$

(d) A distribution consists of 3 groups with frequencies 45, 40 and 65 having their means 2, 2.5 and 2 respectively. Find the mean of the combined distribution. [3]

2.(a) What is dispersion ? What are the common measures of dispersion ? How should a measure of dispersion change when all values of the variable are increased or decreased (a) by the same amount ? (b) In the same proportion ? Judge in this light the different measures of dispersion. [6]

p.t.o.

Contd..... Q.No.2

(b) What is meant by relative dispersion ? Define Coefficient of variation and explain its usefulness.

[6]

(c) The first of two samples has 100 items with mean 15 and s.d. 5. If the whole group has 250 items with mean 15.6 and s.d. $\sqrt{13.44}$. Find the standard deviation of the second group.

[8]

3.(a) What are skewness and Kurtosis ? Give some suitable measures of skewness and Kurtosis.

[3]

(b) Using Cauchy-Schwarz inequality, or otherwise, prove that

(i) $\beta_2 \geq 1$ and (ii) $\beta_2 \geq \beta_1 + 1$

[5]

(c) For a distribution of 25 heights, calculations showed that the mean, standard deviation, β_1, β_2 were 54 inches, 3 inches, 0 and 3 respectively. It was however discovered on checking that the two items 64 and 50 in the original data were wrongly written in place of correct values 62 and 52 inches respectively. Calculate the correct frequency constants.

[12]

4.(a) Show that Poisson distribution is a limiting case of Binomial distribution.

[6]

(b) Prove that for the Binomial distribution, the following relation

$$\mu_{r+1} = pq\left(n r \mu_{r-1} + \frac{d\mu_r}{dp}\right)$$

Contd..... 3/-

Contd..... Q.No.4.(b)

holds, where μ_r is the rth order central moment and p is the parameter of Binomial distribution.

[8]

(c) The Pascal distribution is defined by

$$f(x) = \frac{1}{1+\mu}\left(\frac{\mu}{1+\mu}\right)^x, \quad x = 0, 1, 2, \ldots$$

where $\mu > 0$.

Find the mean and variance of the distribution.

[6]

5. From records of 10 Russian army corps kept over 20 years, the following data was obtained showing the number of deaths caused by the Kicks of a horse. Determine m the average number of deaths per army corps per annum and calculate the theoretical poisson frequencies.

Number of deaths per army corps per annum	0	1	2	3	4	Total
Frequency of occurrence	109	65	22	3	1	200

What is the probability of more than two deaths caused by the Kicks of a horse ?

(18 + 2) = [20]

:bcc: